MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 4, 512–516
https://doi.org/10.1080/23802359.2024.2338560

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

Characterization of the complete chloroplast genome of *Eutrema deltoideum* (Brassicaceae)

Zilu Zhang[a], Guoqian Hao[b] and Zhenhui Kang[a]

[a]School of Bioengineering, Sichuan University of Science & Engineering, Yibin, China; [b]Faculty of Agriculture, Forestry and Food Engineering, Yibin University, Yibin, China

ABSTRACT

Eutrema deltoideum (Hook. f. et Thoms.) has been recognized as a potentially important vegetable and medicinal resource. In this study, we present the complete chloroplast genome of *E. deltoideum* and conduct a phylogenetic analysis. The chloroplast genome is 154,051 bp long and consists of a large single-copy (LSC) region of 84,149 bp, two inverted repeat (IR) regions of 26,065 bp each, and a small single-copy (SSC) region of 17,772 bp. It contains 132 complete genes, including 87 protein-coding genes, 8 ribosomal RNA genes, and 37 tRNA genes. Additionally, we identified 78 simple sequence repeats (SSRs). The phylogenetic tree reveals that *E. deltoideum* is closely related to *E. heterophyllum*, and the Eutrema genus is monophyletic. This study provides valuable information about *E. deltoideum* and enhances our understanding of its taxonomic classification.

ARTICLE HISTORY
Received 8 September 2023
Accepted 31 March 2024

KEYWORDS
Eutrema deltoideum; chloroplast genome; phylogeny

Introduction

The genus *Eutrema* R. Br. is of significant economic importance within the Brassicaceae family, comprising approximately 40 species that are primarily found in East Asia and neighboring regions (Hao et al. 2017; German and Al-Shehbaz 2018). *Eutrema* plants are renowned for their abundant production of isothiocyanates, a class of natural compounds known for their potent pungent flavor and remarkable antimicrobial properties (Fuke et al. 1997; Qiu et al. 2019; Yamane et al. 2023). *E. deltoideum* (Hook. f. et Thoms.) Schulz 1924 is a promising botanical resource with significant potential as both a valuable vegetable and medicinal plant (Niedenzu in Engler 1928; Hao et al. 2017). As a wild plant, *E. deltoideum* exhibits remarkable adaptability to extreme environments and thrives at high altitudes ranging from 3000 to 4000 m in the western regions of China. However, there is currently a significant dearth of molecular data regarding *E. deltoideum*. In this study, we aimed to address this gap by conducting a comprehensive analysis of the complete chloroplast (cp) genome sequences of *E. deltoideum*. By characterizing and examining the entirety of its cp genome sequence, we sought to establish a valuable resource that can be utilized for further exploration of the evolutionary relationships within the Brassicaceae family.

Materials and methods

Plant material, DNA extraction, and sequencing

Fresh leaves of *E. deltoideum* were collected from Lasa, Tibet, China; coordinates 30°15′ N, 91°29′ E, 4390 m (Figure 1). A voucher specimen was identified by Zhenhui Kang and deposited at the herbarium, College of Life Sciences, Sichuan University (SZ), with the voucher number ED20230304 and under the charge of Zhenhui Kang (150363146@qq.com). Total genomic DNA was extracted using the modified CTAB method (Doyle and Doyle 1987). The high-quality DNA was sequenced on the Illumina HiSeq 2000 platform, resulting in approximately 10 GB of paired-end clean reads. The complete chloroplast (cp) genome was assembled using SPAdes v3.15.4 (Bankevich et al. 2012), with *Eutrema yunnanense* (KT270357.1) as the reference sequence, by utilizing the clean reads obtained and executing the command: spades.py −1 1.fq −2 2.fq -o results -k 21,77,127. The annotation of the cp genome was performed using CPGAVAS2 (Shi et al. 2019). After the initial annotation, manual adjustments were made to further refine the results. Finally, Chloroplot was used to visualize the cp genome (Zheng et al. 2020).

Repeats and SSR analysis

Simple sequence repeats (SSRs) were identified using MISA, including mono-, di-, tri-, tetra-, penta-, and hexa-nucleotides with minimum numbers of 10, 6, 5, 4, 4, and 4, respectively (Beier et al. 2017). In addition, REPuter was used to calculate palindromic repeats, forward repeats, reverse repeats, and complementary repeats with the following parameters: hamming distance and minimum repeat size were 3 and 30, respectively (Kurtz et al. 2001).

CONTACT Zhenhui Kang 150363146@qq.com School of Bioengineering, Sichuan University of Science & Engineering, Yibin, Sichuan, China.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2338560.

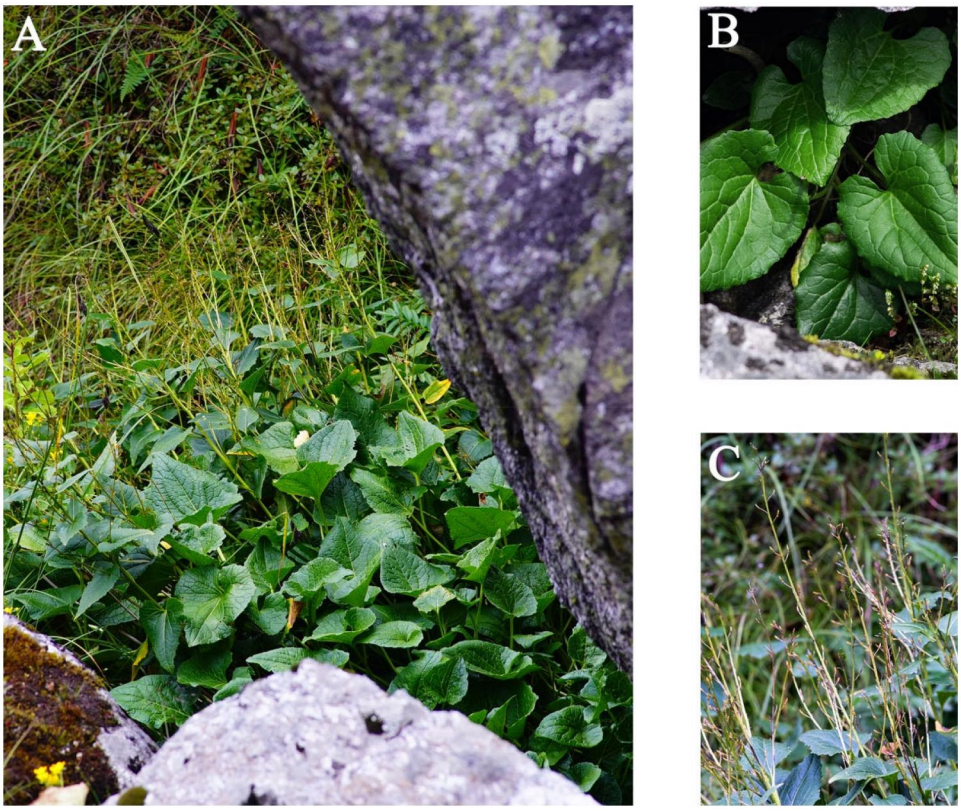

Figure 1. The reference image of *E. deltoideum*. (A) *E. deltoideum*, (B) Leaves, (C) Fruit. *E. deltoideum* is morphologically most similar to *E. baimashanicum*, from which it is distinguished by having smaller flowers, curved, oblong to ovate fruit. This image was taken by Hao at coordinates 27°29′05″ N 88°54′25″ E.

Phylogenetic analysis

To further clarify the phylogenetic position of *E. deltoideum*, the cp genomes of 18 representative Brassicaceae sequences were obtained from NCBI to construct the phylogeny. The dataset included 11 sequences from *Eutrema*, 3 sequences from *Isatis*, and 2 sequences from *Arabis*, 2 sequences from *Descurainia*, and *Arabidopsis thaliana*. After adjustment and alignment with MAFFT v7.49 (Katoh and Standley 2013), a maximum-likelihood (ML) phylogenetic tree was constructed using RAxML v8.2.12 (Stamatakis 2014) with the GTR GAMMA model and 1000 bootstrap replicates. *Arabidopsis thaliana*, *Descurainia sophia*, and *Descurainia erodiifolia* were used as outgroups for analyses of the phylogenetic relationships.

Results

General features of the cp genome

The *E. deltoideum* cp genome is 154,051 bp in length, consisting of a typical quadripartite structure with a large single-copy (LSC) region of 84,149 bp, two inverted repeat (IR) regions of 26,065 bp each, and a small single-copy (SSC) region of 17,772 bp. The entire cp genome contains 132 complete genes, including 87 protein-coding genes (87 PCGs), 8 ribosomal RNA genes (8 rRNAs), and 37 tRNA genes (37 tRNAs). Most genes occur in a single copy, while 19 genes are duplicated, including 4 rRNAs (*rrn*4.5, *rrn*5, *rrn*16, and *rrn*23), 7 tRNAs (*trn*N-GUU, *trn*R-ACG, *trn*A-UGC, *trn*E-UUC, *trn*V-GAC, *trn*L-CAA, and *trn*I-CAU), and 8 PCGs (*rps*7, *rpl*2, *rpl*23, *ndh*B, *ycf*2, *ycf*15, *ycf*1, and *rps*12). The overall GC content of the cp genome is 36.35%, while the corresponding values of the LSC, SSC, and IR regions are 34.08%, 29.32%, and 42.43%, respectively. (Figure 2; Figure S1; Figure S2). Then the annotated cp genome was submitted to GenBank with the accession number OQ473639.

Repeats analysis

In the present study, a total of 78 simple sequence repeats (SSRs) were identified in the cp genome of *E. deltoideum*. These SSRs consisted of 73 mononucleotide repeats, 5 dinucleotide repeats, and no trinucleotide, tetranucleotide, pentanucleotide, or hexanucleotide repeats were detected. Most of the SSRs were mononucleotides, accounting for 94% of the total. In the cp genome of *E. deltoideum*, SSRs were most abundant in the LSC region and lowest abundance in the IR region. In addition, we detected 40 long repeats, including 18 forward repeats and 22 palindromic repeats (Table S1).

Phylogenetic analysis

To understand the phylogenetic position of *E. deltoideum* within the Eutrema genus. This analysis yielded a phylogenetic tree with 100% bootstrap support for all its nodes, showing the reliability of the phylogeny (Figure 3). The phylogenetic tree shows that *E. deltoideum* and *E. heterophyllum* were closely related, and all species of *Eutrema* comprised a monophyletic group within the Brassicaceae.

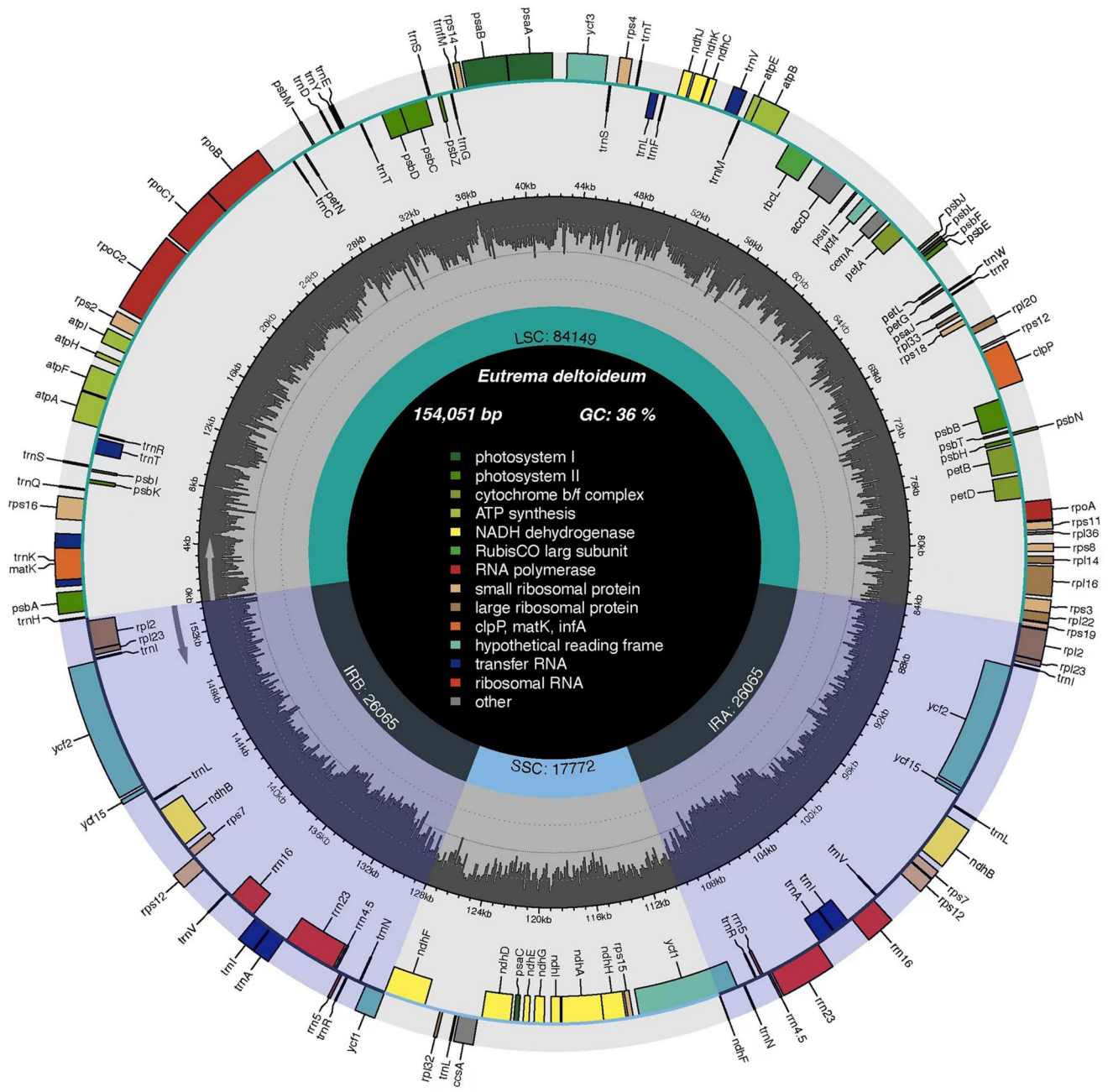

Figure 2. Circular map of the cp genome of *Eutrema deltoideum*. The direction of transcription is indicated by the orientation of genes, with clockwise transcription indicated for genes on the inside of the circle and counter-clockwise transcription indicated for genes on the outside of the circle. The inner circle shows the distribution of at (light grey) and GC (dark grey) content across the genome.

Discussion and conclusion

In this study, the entire cp genome of *E. deltoideum* was comprehensively sequenced, assembled and annotated, which contains 132 complete genes and 78 SSRs.

The structure and sequence of chloroplast harbor rich information, which is importance in elucidating the origins, evolution, and the phylogenetic relationships among different species. We believe that the discovered cp genome will serve as a valuable resource for advancing the phylogenetic research on *Eutrema*.

A total of 19 cp genome sequences from Brassicaceae species were aligned for the phylogenetic analysis, in which *E. deltoideum* and *E. heterophyllum* were clustered into one group, showing their close evolutionary relationship. However, *E. baimashanicum* is morphologically very similar to *E. deltoideum*, which is absent from the phylogenetic tree. To attain more accurate phylogenetic relationships, additional genome of *Eutrema* need to be sequenced.

Ethical approval

The sample (*E. deltoideum*) was collected strictly by guidelines provided by the Sichuan University and Chinese regulations. No ethical approval/permission is required in this study.

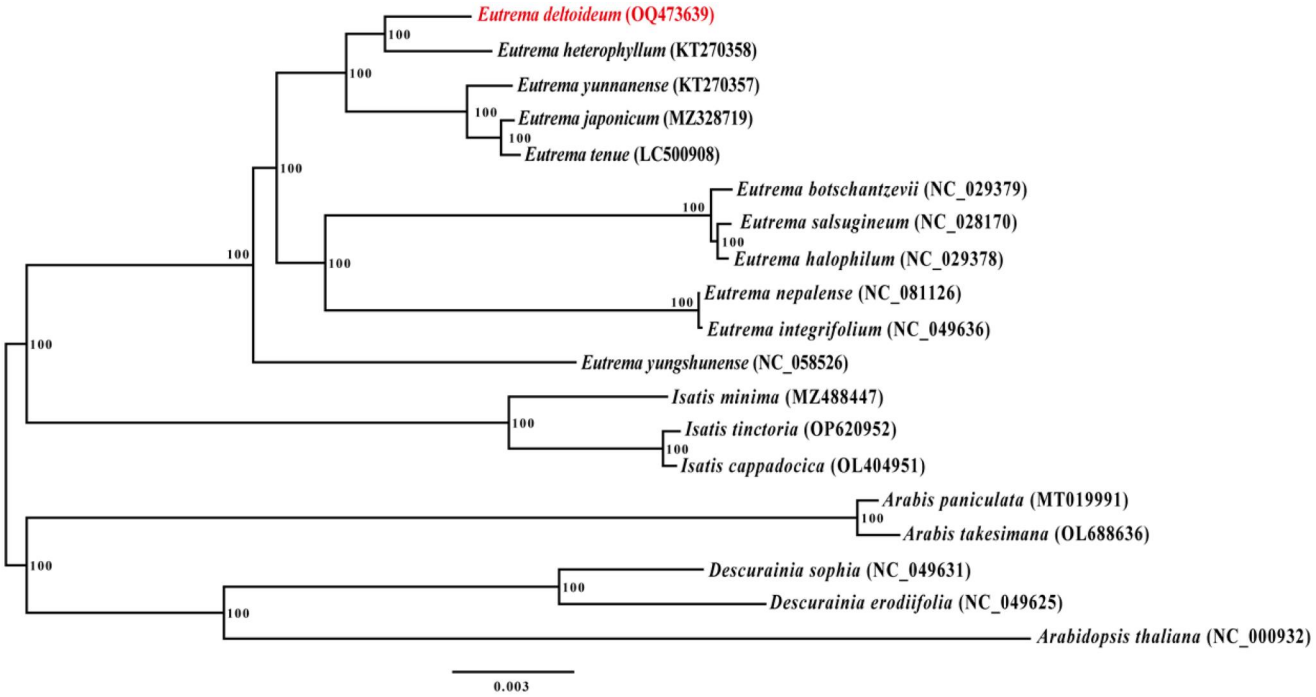

Figure 3. The maximum-likelihood phylogenetic tree was constructed based on the 19 complete cp genome sequences. Numbers in each node indicated the bootstrap support values. The following sequences downloaded from NCBI were used: *E. japonicum* (MZ328719.1) (Liu et al. 2022), *E. tenue* (LC500908.1) (Haga et al. 2019), *E. yunnanense* (KT270357.1) (Hao et al. 2017), *E. heterophyllum* (KT270358.1) (Hao et al. 2017), *E. integrifolium* (NC_049636.1) (Chai et al. 2022), *E. botschantzevii* (NC_029379.1), *E. halophilum* (NC_029378.1), *E. nepalense* (NC_081126.1), *E. salsugineum* (NC_028170.1), *E. yungshunense* (NC_058526.1), *I. cappadocica* (OL404951.1) (Fang et al. 2022), *I. tinctoria* (OP620952.1) (Su et al. 2023), *I. minima* (MZ488447.1), *A. paniculata* (MT019991.1) (Wang et al. 2022), *A. takesimana* (OL688636.1) (Kim et al. 2023), *D. erodiifolia* (NC_049625.1) (Mabry et al. 2020), *D. sophia* (NC_049631.1) (Chen et al. 2016), *A. thaliana*, (NC_050648) (Park et al. 2020).

Authors' contributions

Zhenhui Kang was responsible for the experiment design and article revision. Zilu Zhang collected the material and drafted the manuscript. Guoqian Hao assembled and annotated the cp genome sequences. All authors approved the final version of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the Project of the Department of Sichuan Science and Technology (2020YFN0023)

Data availability statement

The genome sequence data of this study are openly available in NCBI under the accession number OQ473639. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA937053, SRR23562289, and SAMN33388025, respectively.

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477. doi:10.1089/cmb.2012.0021.

Beier S, Thiel T, Münch T, Scholz U, Mascher M., 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics. 33(16):2583–2585. doi:10.1093/bioinformatics/btx198.

Chai W, Li H, Chen X, Qiu J, Xia P. 2022. The whole chloroplast genome of *Neomartinella yungshunensis* (Brassicaceae), an unusual wild plant. Mitochondrial DNA B Resour. 7(6):1096–1098. doi:10.1080/23802359.2022.2087543.

Chen H, Deng T, Yue J, Al-Shehbaz IA, Sun H. 2016. Molecular phylogeny reveals the non-monophyly of tribe Yinshanieae (Brassicaceae) and description of a new tribe, Hilliellelleae. Plant Divers. 38(4):171–182. doi:10.1016/j.pld.2016.04.004.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissues. Phytochem Bull. 19:11–15.

Fang ZX, Qin XJ, Li Q. 2022. The complete chloroplast genome of Isatis cappadocica Desv. (Brassicaceae). Mitochondrial DNA B Resour. 7(5):846–847. doi:10.1080/23802359.2022.2074802.

Fuke Y, Haga Y, Ono H, Nomura T, Ryoyama K., 1997. Anti-carcinogenic activity of 6-methylsulfinylhexyl isothiocyanate-, an active anti-proliferative principal of wasabi (*Eutrema wasabi* Maxim.). Cytotechnology. 25(1-3):197–203. doi:10.1023/A:1007918508115.

German D, Al-Shehbaz I. 2018. One of the last unresolved species in Cochlearia sens. trad., West Himalayan *C. minutissima*, is *Eutrema* (Brassicaceae). Phytotaxa. 338(2):216–218. doi:10.11646/phytotaxa.338.2.7.

Haga N, Kobayashi M, Michiki N, Takano T, Baba F, Kobayashi K, Ohyanagi H, Ohgane J, Yano K, Yamane K. 2019. Complete chloroplast genome sequence and phylogenetic analysis of wasabi (*Eutrema japonicum*) and its relatives. Sci Rep. 9(1):14377. doi:10.1038/s41598-019-49667-z.

Hao G, Al-Shehbaz IA, Ahani H, Liang Q, Mao K, Wang Q, Liu J. 2017. An integrative study of evolutionary diversification of *Eutrema* (Eutremeae, Brassicaceae). Bot J Linnean Soc. 184(2):204–223. doi:10.1093/botlinnean/box024.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/molbev/mst010.

Kim HM, Kim JH, Lee JH, Kim GM, Lee MH, Park CY, Kim DH, Lee DH, Kim KM, Na CS. 2023. Dormancy-release and germination improvement of Korean bellflower (*Campanula takesimana* Nakai), a rare and endemic plant native to the Korean peninsula. PLoS One. 18(10):e0292280. doi:10.1371/journal.pone.0292280.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R., 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 29(22):4633–4642. doi:10.1093/nar/29.22.4633.

Liu H, Qiao Q, Ye X, Guo Y, Ye B, Xu Q, Zhou X, Li B. 2022. The reference genome and organelle genomes of wasabi (*Eutrema japoniacum*). Front Genet. 13:1048264. doi:10.3389/fgene.2022.1048264.

Mabry ME, Brose JM, Blischak PD, Sutherland B, Dismukes WT, Bottoms CA, Edger PP, Washburn JD, An H, Hall JC, et al. 2020. Phylogeny and multiple independent whole-genome duplication events in the Brassicales. Am J Bot. 107(8):1148–1164. doi:10.1002/ajb2.1514. Epub 2020 Aug 24.

Niedenzu in Engler. 1928. Pflanzenreich IV. 141 (Heft 91): 1–246.

Park J, Xi H, Kim Y. 2020. The complete chloroplast genome of *Arabidopsis thaliana* isolated in Korea (Brassicaceae): an investigation of intraspecific variations of the Chloroplast Genome of Korean *A. thaliana*. Int J Genomics. 2020:3236461–3236418. doi:10.1155/2020/3236461.

Qiu L, Zhang M, Bhandari B, Fang Z, Liu Y. 2019. Physicochemical and nutritional properties of wasabi (*Eutrema yunnanense*) dried by four different drying methods. Drying Technol. 37(3):363–372. doi:10.1080/07373937.2018.1458318.

Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 47(W1):W65–W73. doi:10.1093/nar/gkz345.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313. doi:10.1093/bioinformatics/btu033.

Su Y, Zhang M, Guo Q, Wei M, Shi H, Wang T, Han Z, Liu H, Liu C, Huang J. 2023. Classification of *Isatis indigotica* Fortune and *Isatis tinctoria* Linnaeus via comparative analysis of chloroplast genomes. BMC Genomics. 24(1):465. doi:10.1186/s12864-023-09534-8.

Wang H, Gan C, Luo X, Dong C, Zhou S, Xiong Q, Weng Q, Hu X, Du X, Zhu B. 2022. Complete chloroplast genome features of the model heavy metal hyperaccumulator *Arabis paniculata* Franch and its phylogenetic relationships with other Brassicaceae species. Physiol Mol Biol Plants. 28(4):775–789. doi:10.1007/s12298-022-01151-1.

Yamane K, Yamada-Kato T, Haga N, Ishida K, Murayama S, Kobayashi K, Okunishi I. 2023. Allyl isothiocyanate and 6-(methylsulfinyl) hexyl isothiocyanate contents vary among wild and cultivated wasabi (*Eutrema japonium*). Breed Sci. 73(3):237–245. doi:10.1270/jsbbs.22080.

Zheng S, Poczai P, Hyvönen J, Tang J, Amiryousefi A. 2020. Chloroplot: an online program for the versatile plotting of organelle genomes. Front Genet. 11:576124. doi:10.3389/fgene.2020.576124.